Estimation Methodology for Natural Gas Production In the Gulf of Mexico

ASA Committee on Energy Statistics
Energy Information Administration
October 16, 17, 2003
Washington, DC

John Wood, EIA; Gary Long, EIA; and Lucy Luo, Z, Inc.

Estimation Methodology for Natural Gas Production

In the Gulf of Mexico

Introduction

The Gulf of Mexico (GOM) accounts for about 25% of the nation's total natural gas production. The EIA needs good GOM production estimates for several of its programs. However, the first release of Minerals Management Service (MMS) accepted preliminary data (95 percent of wells reported) lags by about 1.5 years. Reliable, final (essentially complete) production data from (MMS) can add another 6 months to a year to the lag. Because of these large lag times, the Reserves and Production Division (RPD) in EIA's Office of Oil and Gas (OOG) began developing and using various methodologies to create GOM natural gas monthly production estimates based on data supplied by the MMS. This document focuses on the methodology currently being used.[1]

The MMS began releasing their raw (unedited, uncertified) well production data in February 2003. After some editing, this data allows RPD/OOG to construct well production distributions of the early but incomplete well production data. The current method uses the early reported well data distribution as a sample of the final complete distribution to produce estimates of final monthly production data. These estimates for a given month become more reliable over time as the well data approaches completeness. A monthly estimate can be verified at the well level by EIA when the well data are complete. As in any real-world estimation process, all of the methods may require expert judgment or analyst override, especially when unanticipated phenomena such as hurricanes occur in the GOM.

Data Description and Preparation

The process naturally starts with the MMS data and is dependent on their ability to collect and report these data. Starting in February the MMS made its new (suspended) well/completion production data available to EIA and the public on its Website, http://www.gomr.mms.gov/homepg/pubinfo/freeasci/product/freeprod.html. In part, this was in response to EIA's request for these data. The MMS well/completion data is now available in three kinds of files:

1) Accepted or verified data that has 95% or more of well/completions reported (latest month meeting this condition is March 2002),
2) Accepted data that have less than 95% of wells reported (subsequent data), and
3) Suspended data (newly available un-accepted, un-verified, un-edited data).

Data edited by RPD from these three files are combined to yield the total reported production. Having access to these data has permitted the development of the

[1] Previous methodologies have been based on average month-to-month changes in historical GOM production data, average month-to-month changes in Texas production data, a simple linear model based on Texas production data, and the smoothed means of individual well production data.

methodologies to estimate final production described here. All these data are downloaded from the MMS Website in 8 separate files.

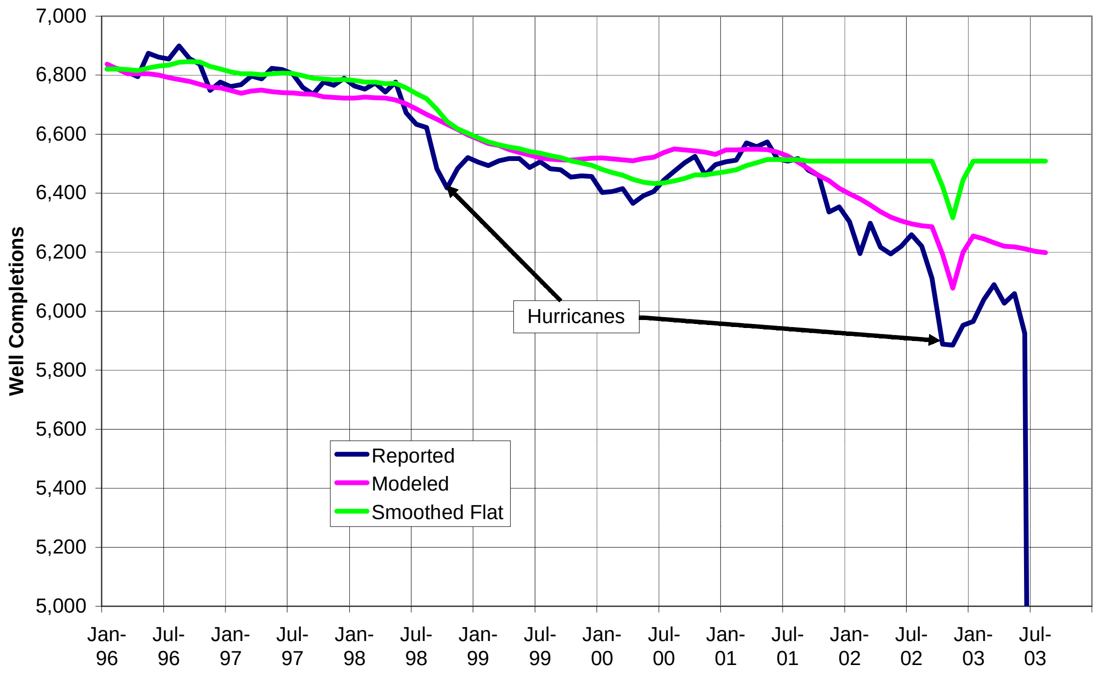
Historical data on the MMS Website go back to 1996 and are updated occasionally (1996 through 2003 last updated September 2003). Current accepted data are updated monthly. Suspended data are currently updated approximately twice per month. Accepted data are in zipped delimited text files, and the suspended data are in zipped Excel files. This detailed production data are by well completion by month.

SAS programs have been written to convert the downloaded delimited data into two SAS data sets of summarized monthly production. One data set contains all accepted data and the other contains the suspended data. Of the roughly 64,530 records in the suspended data (released date: 09-15-2003), about 3,700 were duplicates in the data processed in September 2003. Also, there are about 800 duplicates between the suspended and accepted data. SAS programs are used to identify and delete the duplicate records. There are approximately 6,500 records (completions) each month with gas production greater than zero (these include associated gas from oil completions).

Gas Production Estimation Methodology

The reported data is essentially complete in mid 2001 and progressively less complete closer to current months. Figure 1 shows reported well completions dropping from about 6,500 in September 2001 to 5,900 in June 2003. A requirement of the methodology is to have the expected well completions for each incomplete month. Then the expected completions and average production per completion can be used to estimate production in each month. Along with the reported well completions shown in Figure 1, two other estimates of expected well completions are shown. These will be discussed later.

Figure 1. Well Completions with Gas Production > 0



Simple Model

For any distribution the mean or average production per completion is:

$$M_i = \frac{P_i}{W_i}$$

Where:

M_i = Mean production per completion for month i
P_i = Total production for month i
W_i = Total producing completions for month i

The simplest model for the production is

$$P'_i = W'_i * M'_i$$

Where:

P'_i = Modeled production for month i
W'_i = Modeled or expected number of completions for month i
M'_i = Modeled mean production per completion for month i

However, the mean production per completion is skewed, so an estimation based solely on the mean may not be a good estimate.

Figure 2. Variable: Well_Rate (Month: May 2001)

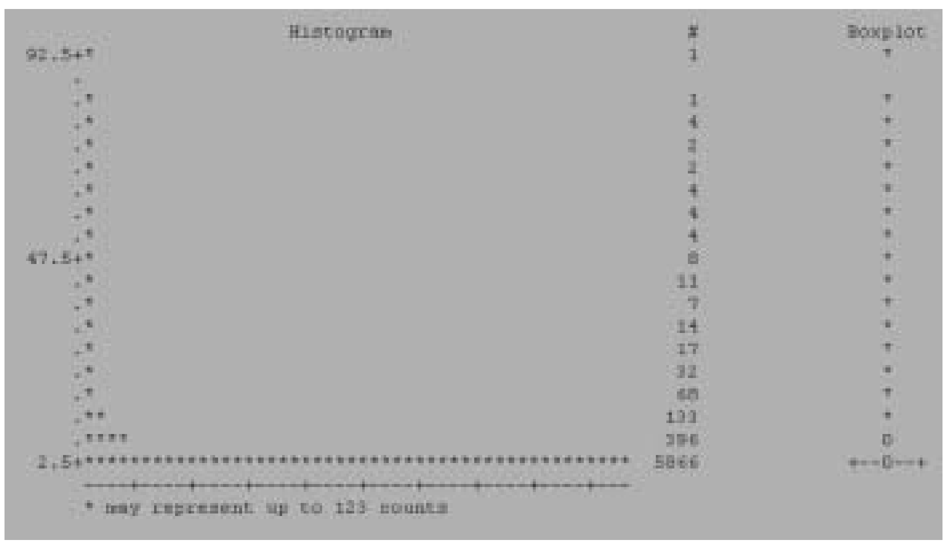
Moments			
N	6574	Sum Weights	6574
Mean	2.16640601	Sum Observations	14241.9531
Std Deviation	5.91859516	Variance	35.0297687
Skewness	6.87008015	Kurtosis	63.735333
Uncorrected SS	261104.523	Corrected SS	230250.67
Coeff Variation	273.198797	Std Error Mean	0.07299679

Basic Statistical Measures			
Location		Variability	
Mean	2.166406	Std Deviation	5.91860
Median	0.318597	Variance	35.02977
Mode	0.000290	Range	91.78294

Tests for Normality				
Test	Statistic		p Value	
Kolmogorov-Smirnov	D	0.357172	Pr > D	<0.0100
Cramer-von Mises	W-Sq	261.0979	Pr > W-Sq	<0.0050
Anderson-Darling	A-Sq	1284.144	Pr > A-Sq	<0.0050

Quantiles (Definition 5)	
Quantile	Estimate
100% Max	91.7830

Quantiles (Definition 5)	
Quantile	Estimate
99%	2.84755
95%	9.70106
90%	5.45597
75% Q3	1.71177
50% Median	0.318597
25% Q1	0.077419
10%	0.0234194
5%	0.0103548
1%	0.00135484
0% Min	0.000032258



12 Classes Model

The data is divided into 12 classes and the latest six months of complete production data (from April 2001 to September 2001) were used to create the expected distribution for each month (Table 1). If production per completion is less than 1.0 MMCF (most of these are oil well completions) or over 100 MMCF (rare but highly productive gas well completions), then they are defined as classes 1 and 12 respectively. As for other classes, the well completion is an exponential distribution (Figure 3.) with the formula of Wij=690(A)^(J-2). EXCEL solver is used to determine the coefficient A. Then based on the calculated well completions in the six-month calibration or expected data distribution, the class boundaries are determined for each class (Table 1).

Figure 3. Well Distribution by Class

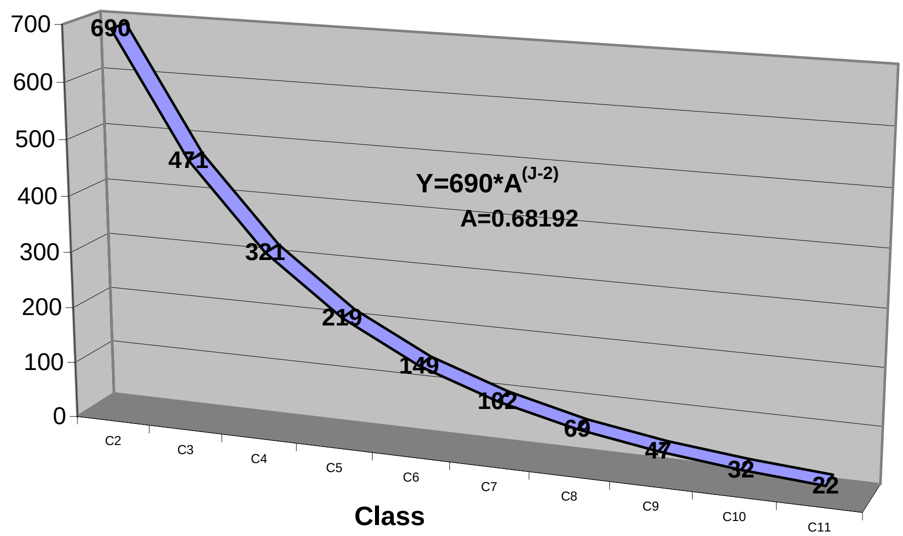


Table1. Class Determination Method

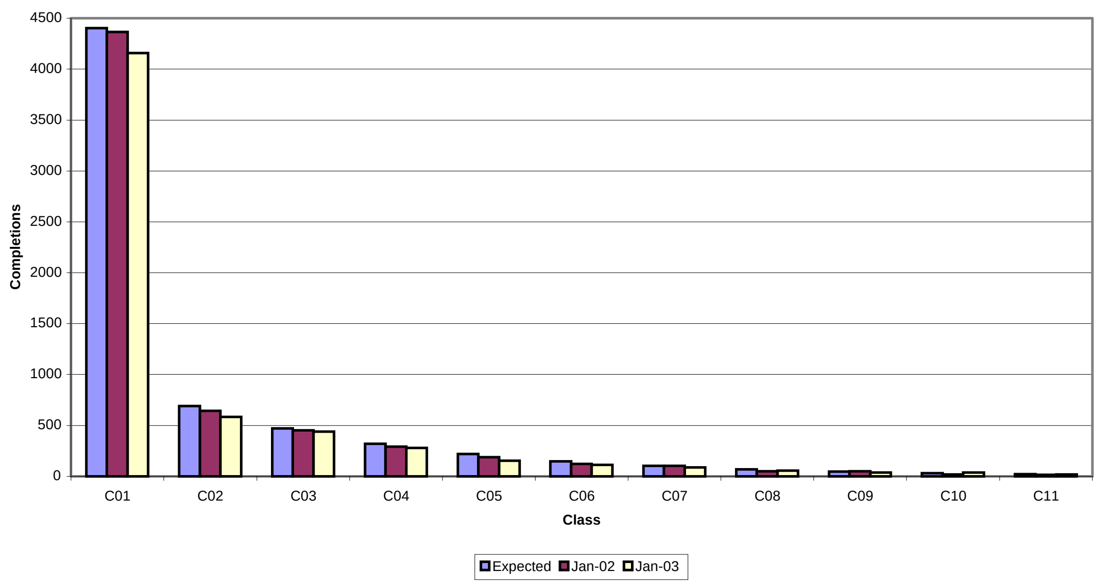
Class(J)	Well Completions Number	Formula	Well Production Rate (MMCF/Day)
1	4402		0<P<1.0
2	690	Wij=690(0.68192)^(j-2)	1.0<=P<2.078
3	471	Wij=690(0.68192)^(j-2)	2.078<=P<3.601
4	321	Wij=690(0.68192)^(j-2)	3.601<=P<5.491
5	219	Wij=690(0.68192)^(j-2)	5.491<=P<7.951
6	149	Wij=690(0.68192)^(j-2)	7.951<=P<10.945
7	102	Wij=690(0.68192)^(j-2)	10.945<=P<15.15
8	69	Wij=690(0.68192)^(j-2)	15.15<=P<20.995
9	47	Wij=690(0.68192)^(j-2)	20.995<=P<31
10	32	Wij=690(0.68192)^(j-2)	31<=P<50
11	22	Wij=690(0.68192)^(j-2)	50<=P<100
12	Uncertain		100<=P

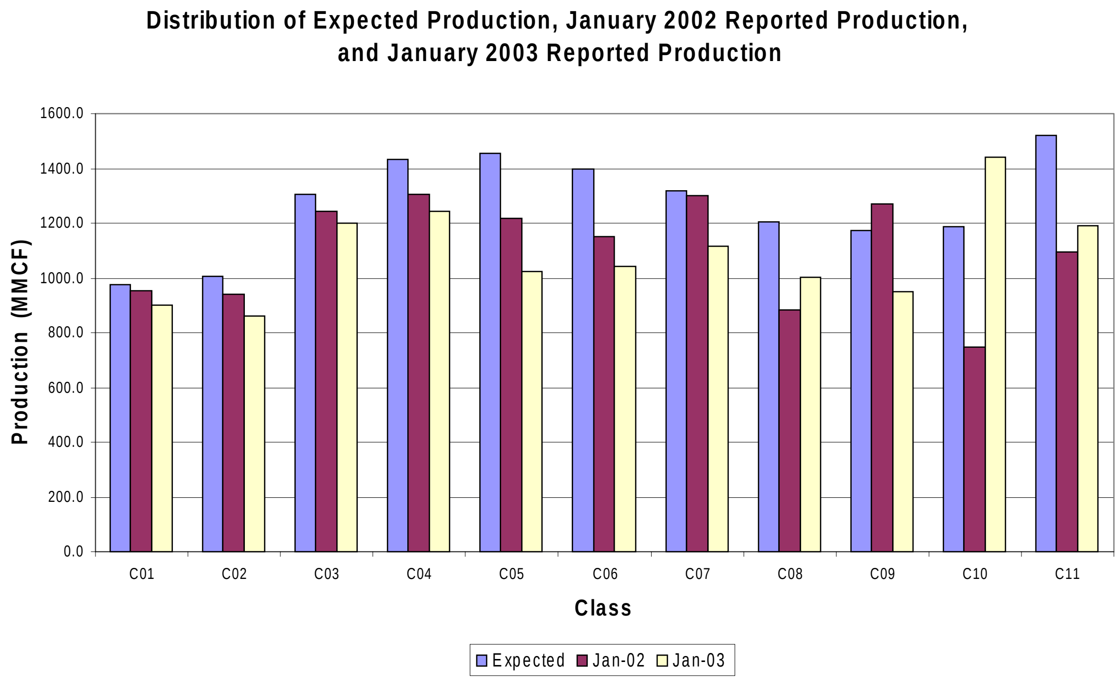
The basic concept assumes that incomplete data in recent months is a sample distribution of what will ultimately be reported as the final distribution. The expected distribution determines the expected completions for each incomplete class. If the wells for a given month equal or exceed the expected number of wells, then that month is accepted as essentially complete. For months with fewer wells than expected, any class with completions numbering more than the expected distribution is considered complete and accepted as is. For all incomplete classes, the total number of missing completions is allocated to the incomplete classes proportional to the expected number of completions of all classes that are not full. Then, for each class, the product of the number of completions and the reported average production per completion is the estimate of production for that class.

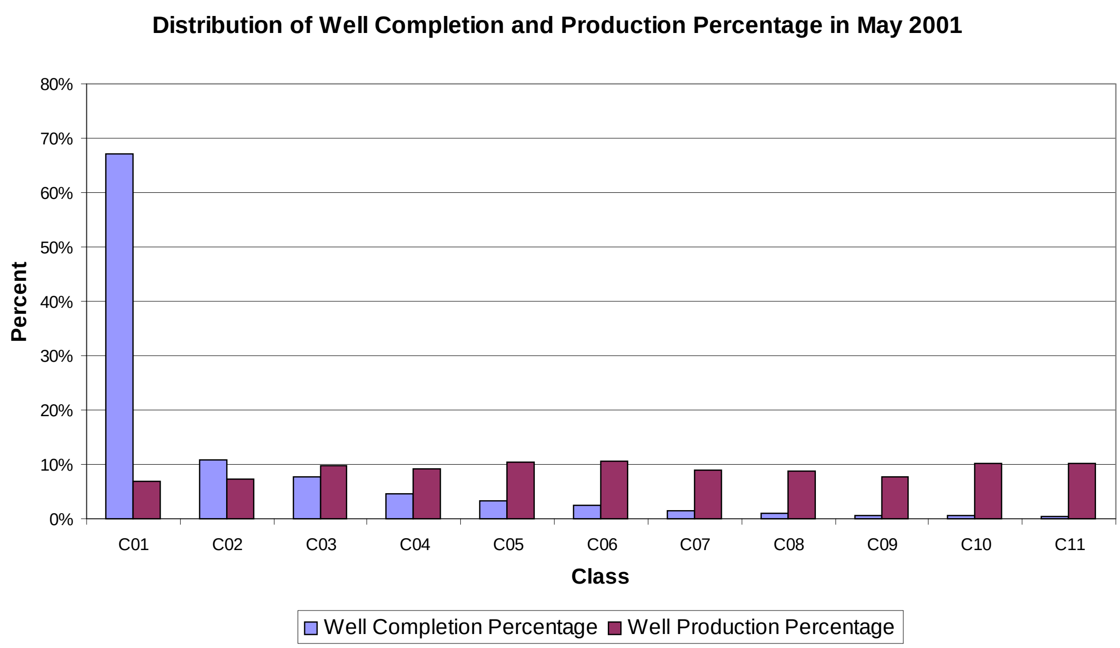
Classes 11 and 12 (production rate greater than 50 MMcf/d) are treated differently. For class 11 we examine each well's historical production record and expert judgment is used to determine where monthly production is missing. For class 12, where each well can change the GOM monthly production by about 1 percent, we use class 12 well completions as reported.

The following chart shows a comparison of two incomplete months to the expected distribution. Note that while not complete the January 2002 distribution is more complete than the January 2003 distribution.

Figure 4. Distribution of the Expected Completions, January 2002 Reported Completions, and January 2003 Reported Completions







Estimates for Class 11

Since class 11 (> 50 MMcf/day) shows significant growth from 1996 to 2002 it requires careful consideration. In 1996 this class had only 3 or 4 wells representing roughly 2 percent of GOM production. In the calibration period, April through September 2001, the class holds about 22 wells with about 10 percent of the production (Figure 5). Table 2 shows the historical production of some of these wells with several months of missing production. As an example, beginning with the August 2003 data, a petroleum engineer estimates four more wells will likely be reported for Class 11 in June 2003, three in May 2003, one each in February through April 2003,two for January 2003, and one in March 2002 (cells highlighted in yellow). Buy the second update in September 2003, all but one well in January 2003 have been reported.

Figure 5. Class 11 Well Count
Large Wells over 50 MMcf/day

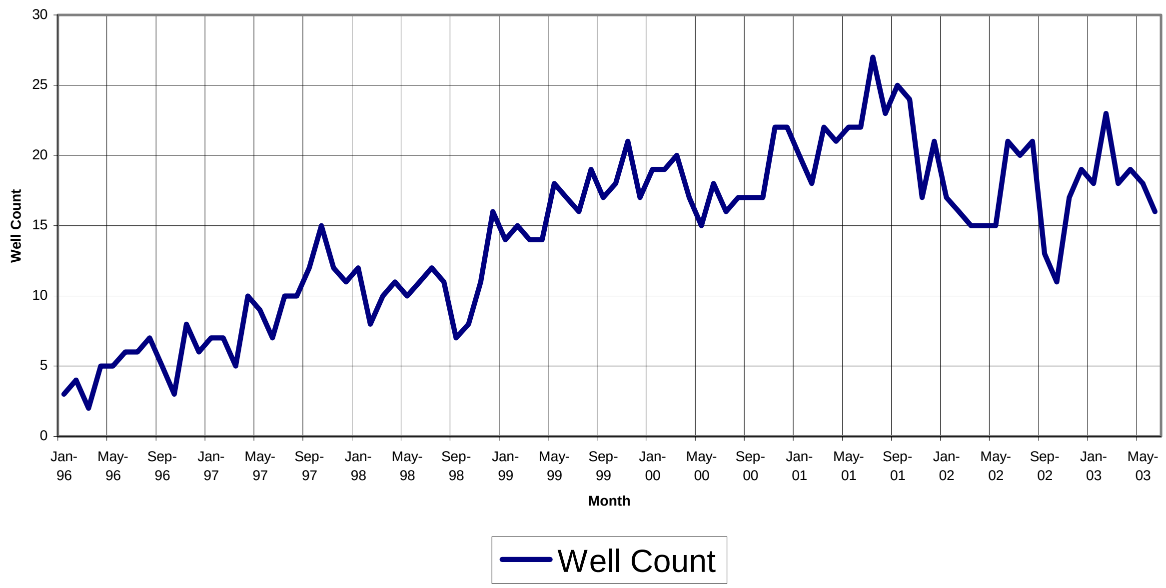


Table 2. Well Records Showing missing reported Value

Aug	Jun 03	May 03	Apr 03	Mar 03	Feb 03	Jan 03	Dec 02	Nov 02	Oct 02	Sep 02	Aug 02	Jul 02	Jun 02	May 02	Apr 02	Mar 02	Feb 02	Jan 02
177154108100	.	.	.	.	.	.	52.6	56.5	51.8	50.9	59.1	65.3	68.9	71.2	67.9	78.7	82.3	92.2
608044022101	.	52.2	52.1	55.2	54.2	.	60.9	60.1	56.9	56.6	59.7	57.6	.	.	.	.	.	.
608044022400	.	47.6	45.9	48.1	47.5	.	51.7	49.6	46.7	47.7	51.9	.	.	.	.	.	.	.
608044023400	98.1	99.0	96.7	52.1	.	.	.	.	.	.	.	.	.	.	.	.	.	.
608044023500	92.3	95.3	90.4	34.3	.	.	.	.	.	.	.	.	.	.	.	.	.	.
608164023900	.	.	45.3	45.4	37.4	48.1	50.9	53.7	45.2	43.6	52.0	57.5	57.6	59.8	59.6	56.4	57.8	61.1
608164024302	.	.	54.4	62.4	50.9	64.7	0.7	66.6	58.2	8.7	70.9	71.7	70.8	70.8	70.5	.	62.7	64.1
608164024700	.	.	78.6	73.0	60.2	77.5	79.2	80.7	63.9	69.0	72.2	80.2	79.0	76.9	51.3	30.9	35.7	38.6
608234000200	.	81.1	84.0	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Sep	Jun 03	May 03	Apr 03	Mar 03	Feb 03	Jan 03	Dec 02	Nov 02	Oct 02	Sep 02	Aug 02	Jul 02	Jun 02	May 02	Apr 02	Mar 02	Feb 02	Jan 02
177154108100							52.6	56.5	51.8	50.9	59.1	65.3	68.9	71.2	67.9	78.7	82.3	92.2
608044022101	50.2	52.2	52.1	55.2	54.2		60.9	60.1	56.9	56.6	59.7							
608044022400	46.4	47.6	45.9	48.1	47.5		51.7	49.6	46.7	47.7	51.9							
608044023400	98.1	99.0	96.7	52.1														
608044023500	92.3	95.3	90.4	34.3														
608164023900		51.2	45.3	45.4	37.4	48.1	50.9	53.7	45.2	43.6	52.0	57.5	57.6	59.8	59.6	56.4	57.8	61.1
608164024302		59.6	54.4	61.8	50.4	64.0	0.7	65.9	58.2	8.7	70.9	71.7	70.8	141.8			62.7	64.1
608164024700		75.4	78.6	73.0	60.2	77.5	79.2	80.7	63.9	69.0	72.2	80.2	79.0	76.9	51.3	30.9	35.7	38.6
608234000200	67.4	81.1	84.0	77.9	79.3	80.4	55.6	51.3	13.1									

Sep-Update	Jun 03	May 03	Apr 03	Mar 03	Feb 03	Jan 03	Dec 02	Nov 02	Oct 02	Sep 02	Aug 02	Jul 02	Jun 02	May 02	Apr 02	Mar 02	Feb 02	Jan 02
177154108100	53.2	55.1	50.7	50.9	53.8	56.1	52.6	56.5	51.8	50.9	59.1	65.3	68.9	71.2	67.9	78.7	82.3	92.2
608044022101	50.2	52.2	52.1	55.2	54.2	.	60.9	60.1	56.9	56.6	59.7	.	.	.	.	.	.	.
608044022400	46.4	47.6	45.9	48.1	47.5	.	51.7	49.6	46.7	47.7	51.9	.	.	.	.	.	.	.
608044023400	98.1	99.0	96.7	52.1	.	.	.	.	.	.	.	.	.	.	.	.	.	.
608044023500	92.3	95.3	90.4	34.3	.	.	.	.	.	.	.	.	.	.	.	.	.	.
608164023900	50.1	51.2	45.3	45.4	37.4	48.1	50.9	53.7	45.2	43.6	52.0	57.5	57.6	59.8	59.6	56.4	57.8	61.1
608164024302	58.7	59.6	57.7	61.8	50.4	64.7	65.7	65.9	58.2	8.7	70.9	71.7	70.8	70.9	79.6	.	62.7	64.1
608164024700	75.8	75.4	78.6	73.0	60.2	77.5	79.2	80.7	63.9	69.0	72.2	80.2	79.0	76.9	51.3	30.9	35.7	38.6
608234000200	67.4	81.1	84.0	77.9	79.3	80.4	55.6	51.3	13.1	.	.	.	.	.	.	.	.	.

The following are the formulas for the twelve class model.

For a distribution with 12 classes:

$$P_i = \sum_{j=1}^{12} P_{i,j}$$

Where j is a class from 1 to 12.

The mean production per completion for any class is given by the following:

$$M_{i,j} = \frac{P_{i,j}}{W_{i,j}}$$

Where

$M_{i,j}$ = Mean production per completion of class j for month i
$W_{i,j}$ = Number of completions of class j for month i

Therefore, the production model for a class, and total are:

$$P'_{i,j} = W'_{i,j} * M_{i,j}$$

$$P'_i = \sum_{j=1}^{12} P'_{i,j}$$

Where $M_{i,j}$ is the actual reported mean production per completion of the sample distribution.

Production is then estimated by the following equation.

$$P'_i = \sum_{j=1}^{10} \left[M_{i,j} * WE'_{i,j} \right] + M_{i,11} * WE'_{i,11} + M_{i,12} * W_{i,12}$$

Where:

P'_i = Modeled production for month i
$M_{i,j}$ = Reported mean production per completion in class j for month i
$WE'_{i,j}$ = Modeled or expected number of completions in class j for month i
$WE'_{i,11}$ = Professional Expected number of completions in class 11 for month i
$W_{i,12}$ = Reported number of completions in class 12 for month i

We have the following options for estimating production:

1) Use the Expected/standard mean production per completion for each class.
2) Use the Reported mean production per completion for each class.
3) Use a Smoothed mean production per completion for each class.
4) Use a Flat Expected well completion count.
5) Use a Modeled Expected well completion count.

Modeled Well Completions from Rig counts

For "normal" months a model based on the rig counts in the GOM can be used to estimate the expected number of wells. The model is calibrated to the six-month reference period (April to September 2001) and supplies an expected number of well completions for each month. The expected completion model is as follows.

$$WE_i = A * e^{\left[\left(\frac{B}{A}\right) * t\right]} + C * SmRigs_i$$

Where:

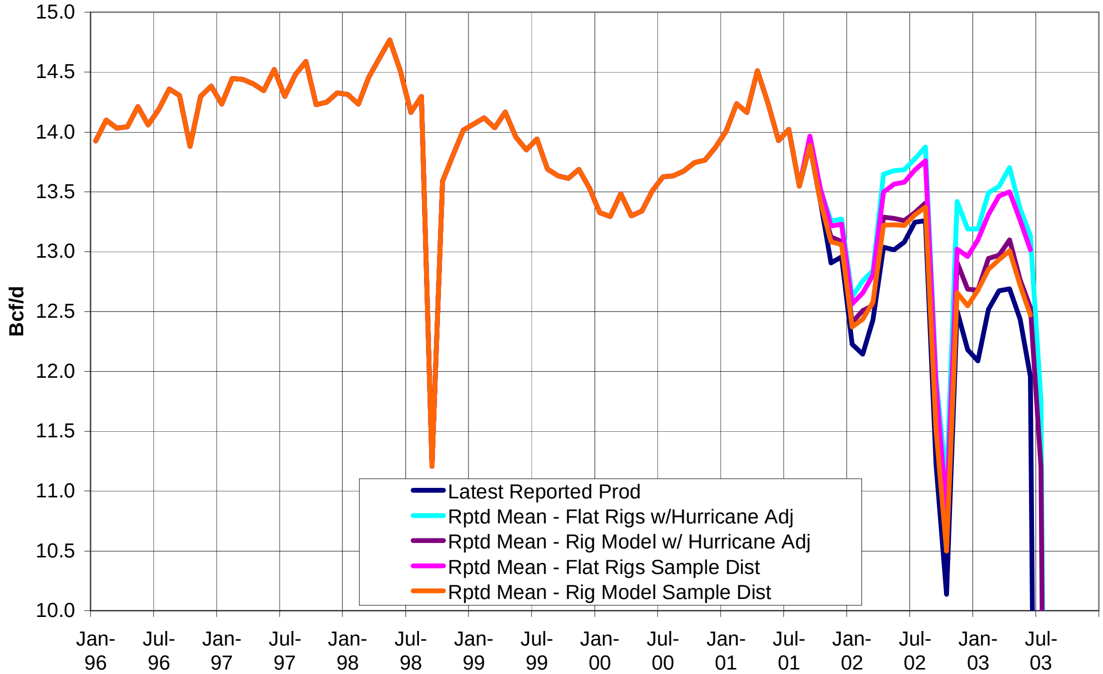
WE_i	= Modeled or expected number of completions for month i
$SmRigs_i$	= Smooth GOM rig count (6 month exponentially smoothed) for month i
A, B, C	= Fit parameters

The first half of the equation is a decline function that reduces the number of completions each month. The second half of the equation adds completions based on the smooth rig count. The resulting expected completions are shown in Figure 1.

Following is a plot (Figure 6.) of the reported production and estimated final production. Reported production is essentially complete through September 2001 (accepted and edited data 99.5 percent complete plus RPD edited suspended data). Major hurricanes or storms occurred in the fall of 1998 and 2002. The estimated production shown here includes an empirical adjustment to the number of completions during the storms of 2002.

Figure 6 shows 4 estimates using the modeled well completions and flat well completions for both a 12 class distribution and a single class distribution. All four cases include an empirical hurricane adjustment. Reported production is essentially complete through September 2001.

Figure 6. GOM Gas Production Estimates Compared to Reported Gas Production



Well Distribution Stability Test

A key part of the procedure is the determination of the expected/standard well distribution based on six months of essentially complete data. We examined the stability of the well distributions using a Chi-Square goodness of fit test. After applying this test to the distribution in the 12 months prior to the 6 standard months and later incomplete months, it was determined that all of the months have a similar distribution (an exception will be discussed later). As an example the following tables show the Chi-Square test for March 2001 and April 2003.

CHI-SQUARE Goodness of fit Test

ProdDate=200103			
Frequency Count			
COUNT	**Frequency**	**Percent**	**Test Percent**
4429	4429	67.64	67.70
701	701	10.71	10.61
472	472	7.21	7.24

ProdDate=200103			
Frequency Count			
COUNT	Frequency	Percent	Test Percent
329	329	5.02	4.94
213	213	3.25	3.37
131	131	2.00	2.29
121	121	1.85	1.57
74	74	1.13	1.06
45	45	0.69	0.72
33	33	0.50	0.49

Chi-Square Test for Specified Proportions	
Chi-Square	6.5276
DF	9
Pr > ChiSq	0.6862
Sample Size = 6548	

ProdDate=200304			
Frequency Count			
COUNT	Frequency	Percent	Test Percent
4175	4175	69.53	67.70
631	631	10.51	10.61

ProdDate=200304			
Frequency Count			
COUNT	Frequency	Percent	Test Percent
405	405	6.74	7.24
273	273	4.55	4.94
170	170	2.83	3.37
129	129	2.15	2.29
90	90	1.50	1.57
57	57	0.95	1.06
48	48	0.80	0.72
27	27	0.45	0.49

Chi-Square Test for Specified Proportions	
Chi-Square	14.2235
DF	9
Pr > ChiSq	0.1146
Sample Size = 6005	

Hurricane Exceptions

For the months of September, October, November, and December 2002, the Chi-Square test indicated that these distributions are different from the standard 6-month distribution. Tropical storm Isidore and Hurricane Lili in September and October affected production in these four months. When wells are shut in for a partial month or even several months, the distribution changes. An empirical downward adjustment to the number of expected completions is necessary for months with a major storm. The Chi-Square test for October 2002 indicates that the distribution is different.

ProdDate=200210			
Frequency Count			
COUNT	Frequency	Percent	Test Percent
4304	4304	73.25	67.70
541	541	9.21	10.61
380	380	6.47	7.24
232	232	3.95	4.94
136	136	2.31	3.37
98	98	1.67	2.29
75	75	1.28	1.57
46	46	0.78	1.06
33	33	0.56	0.72
31	31	0.53	0.49

Chi-Square Test for Specified Proportions	
Chi-Square	93.2510
DF	9
Pr > ChiSq	<.0001
Sample Size = 5876	